

BHAVIN JAWADE

3296 Main Street, Apt C1 • Buffalo, NY, 14214 • +1 716-495-9321 • bhavinjawade@gmail.com
LinkedIn - www.linkedin.com/in/bhavinjawade • **Website** - <https://bhavinjawade.github.io> ·
Github - github.com/bhavinjawade • **Google Scholar** - [Bhavin Jawade](#)

Education

University at Buffalo, SUNY PhD. - Computer Science Advisor - Dr. Venu Govindaraju	Dec 2024 GPA: 3.9
University at Buffalo, SUNY MS. - Computer Science Advisor - Dr. Sargur Srihari	May 2021 GPA: 3.8
SGSITS - Shri G.S. Institute of Technology & Science BE. (Honors) - Information Technology	May 2019 GPA: 3.8

Work Experience

Netflix Research | Research Scientist Intern | Los Gatos, CA May 2024 - Present

- Multimodal (audio-visual) contrastive learning for speech to lip-motion alignment and quality estimation.
- Designed novel loss function “RSMS - Ranking Supervised Multi-Similarity loss” to incorporate positive’s ranking prior within the contrastive learning optimization objective.
- Proposed method achieved 11.3% improvement over the baselines in discerning dubbing lip-sync quality on real-world dubs and other audio understanding tasks
- Work under review at ICASSP 2024.

Yahoo Research | Research Scientist Intern | New York City, NY May 2023 - May 2024

- Proposed SCoT, a novel pre-training strategy to improve zero-shot multimodal compositional retrieval and understanding.
- Performed LoRA based Instruction fine-tuning of quantized Falcon-7B, LLama-2, Mistral-8B, Gemma-2B and SOLAR-10B LLM models to generate synthetic textual triplets. Utilized contrastive representations from textual triplets as a proxy supervision.
- Approach outperformed existing methods by ~10.2% on the Fashion-IQ dataset and ~8% on CIRR benchmark in zero-shot retrieval setting while only utilizing image-caption pairs. Qualitative experiments visualizing SCoT’s latents with UnCLIP demonstrate its implicit compositional properties.
- Research paper accepted at WACV 2025 (accepted in Round 1 - 12% acceptance rate)

Adobe Research | Research Scientist Intern | San Jose, CA May 2022 - Nov 2022

- Proposed a Large-Scale Vision-Language (multimodal foundational models) contrastive pre-training strategy to learn embeddings that generalize over out-of-distribution tasks.
- Introduced a novel loss formulation incorporating tags, metadata, and clauses to create a semantic hierarchy, enhancing the alignment of vision and language embeddings.
- Performed large-scale distributed pre-training using torch-distributed, web-datasets and gradient-checkpointing. Trained on over 100 million image-text pairs on 16 A100 GPUs.
- Method outperforms Open-AI’s CLIP by 6.4% on zero-shot ImageNet retrieval (Top-1) when trained on 50M image-text pairs, 9.2% when trained on 15 million subset and 8.4% when trained on 5 million sub-set

Center for Unified Biometrics and Sensors | Research Assistant | Buffalo, NY Jun 2020 - Present

- IARPA BRIAR Grant | Proposed a uni-modal feature fusion strategies - CoNAN and ProxyFusion
CoNAN received **Best Paper Award, IEEE, IJCB 2023, Slovenia**, and ProxyFusion is accepted at NeurIPS 2024
- Computer vision and vision-language models for retrieval, grounding and alignment. Works presented at WACV 2023 and WACV 2024, Hawaii.
- NSF National AI Institute Grant | Working on a novel benchmark and evaluation metric to estimate hallucination in text conditioned image generation diffusion models. [DPO trained](#) a SOLAR-10.7B LLM using LoRA (Ranked #2 on Hugging-face openLLMLeaderboard, 35K+ Downloads).
- NSF CITER Grant | Developed a multi-task learning based CNN network trained with Deep metric losses (Contrastive, AdaCos) along with minutiae loss to match contactless fingerprints captured using smartphone camera against legacy fingerprints.
Won Russell Agrusa Research Innovation Award, UB CSE, 2021. Received Best CSE PhD Poster Award (2022) and Best UB SEAS Poster Award (2023)
- NSF DIBBS Grant | Implemented full-stack machine learning framework called MLToolkit where the user can drag and drop the different elements to create and execute machine learning pipelines.

Persistent Systems | Machine Learning Engineer | Pune, India

Jul 2019 - Dec 2019

- Built OCR based ML application for banks to auto-analyze POS invoices saving ~\$200,000 & 100 hrs/week.
- Developed an Invoice Management tool for a leading Supply Chain firm to automate the supplier-distributor financing process and decrease finance approval time by 43%.

IIT Madras, Department of Computer Science | Research Intern | Remote

Feb 2019 - May 2019

Advised by Dr. Rupesh Nasre, Professor IIT Madras.

- Designed a highly efficient algorithm for computing execution times in hierarchical task scheduling, targeting updates in million-node Directed Acyclic Graphs (DAGs).
- Achieved performance, with the algorithm processing 1,000 updates in a 1,000,000 node DAG in just 0.8 seconds, demonstrating near-linear $O(n)$ complexity

Selected Publications

ProxyFusion: Efficient Face Feature Aggregation Through Sparse Experts <i>Conference on Neural Information Processing Systems (NeurIPS 2024)</i>	May 2024
SCOT: Self-Supervised Contrastive Pre-training For Zero-Shot Compositional Retrieval <i>IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2025)</i>	Nov 2023
NAPReg: Nouns as Proxies Regularization for Semantically Aware Cross-Modal Embeddings <i>IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2023) (Oral)</i>	Jan 2023
CoNAN: Conditional Neural Aggregation Network For Unconstrained Face Feature Fusion <i>IEEE International Joint Conference on Biometrics (IJCB 2023) (Best Paper Award, Oral)</i>	Sep 2023
Hear The Flow: Optical Flow-Based Self-Supervised Visual Sound Source Localization <i>IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2023) (Oral)</i>	Aug 2022
Audio-visual Representation Learning for Lip-Sync Estimation Through Ranking Augmented Contrastive Training <i>Under Review International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2024)</i>	Sep 2024
RealCQA: Scientific Chart Question Answering as a Test-bed for First-Order Logic <i>IEEE/IAPR International Conference on Document Analysis and Recognition (ICDAR 2023)</i>	Feb 2023
DIOR: Dataset for Indoor-Outdoor Reidentification-Long Range 3D/2D Skeleton Gait Collection Pipeline, Semi-Automated Gait Keypoint Labeling and Baseline Evaluation Methods <i>IEEE International Joint Conference on Biometrics (IJCB 2024)</i>	Aug 2023

Liveness Detection Competition-Noncontact-based Fingerprint Algorithms and Systems IEEE International Joint Conference on Biometrics (IJCB 2023)	Sep 2023
GestSpoof: Gesture Based Spatio-Temporal Representation Learning For Robust Fingerprint Presentation Attack Detection. IEEE International Conference on Automatic Face and Gesture Recognition (FG)	Sep 2023
Conditional Neural Aggregation Network For Unconstrained Long Range Person Feature Fusion IEEE transactions on biometrics, behavior, and identity science (TBIOM) (Invited Journal Paper)	Nov 2023
Deep Metric Learning for Computer Vision: A Brief Overview Book Chapter in Handbook of Statistics, Special Issue - Deep Learning (Invited Book Chapter)	Dec 2022
RidgeBase: A Cross-Sensor Multi-Finger Contactless Fingerprint Dataset IEEE/IAPR International Joint Conference on Biometrics (IJCB 2022)	Jun 2022
Attribute De-biased Vision Transformer (AD-ViT) for Long-Term Person Re-identification IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS 2022)	Sep 2022
Multi Loss Fusion for Matching Smartphone Captured Contactless Finger Images IEEE International Workshop on Information Forensics and Security (WIFS 2021) (Oral)	Sep 2021
Low computation in-device geofencing using hierarchy-based searching for offline usage IEEE International Conference on Inventive Computation Technologies' (ICICT 18') (Oral)	Nov 2018

Research Datasets

RealCQA , Chart Visual Question Answering Dataset	2022
RidgeBase , Contactless Fingerprint Dataset	2021
DIOR , Long-Distance Person ReID Dataset	2023
LivDet 23 , Contactless Fingerprint Liveness Dataset	2023
GestSpoof , Spatio-Temporal Fingerprint Liveness Detection Dataset	2023

Achievements and Awards

IEEE Best Paper Award , IEEE IJCB	2023
Graduate Leadership Award , University at Buffalo	2023
Russell Agrusa Research Innovation Award , University at Buffalo	2021
Research Award: Best SEAS PhD Poster Award , University at Buffalo	2022
Research Award: CSE PhD Poster Award , University at Buffalo	2022
Hackathon Winner , Adobe Code Jam	2022
Best Social AI Idea Award : Blackstone Launchpad	2021
Hackathon Winner , Maple Ridge Hackathon (Govt of British Columbia)	2021
Hackathon Winner , WittyHacks Hackathon	2018
Project Award , Hon' Minister of State (Education Minister, Govt of M.P., India)	2018

Service

Reviewer , Conference on Neural Information Processing Systems, 2024 (NeurIPS)
Reviewer , IEEE CVF Computer Vision and Pattern Recognition Conference, 2023 (CVPR)
Reviewer , IEEE CVF Winter Conference on Applications of Computer Vision, 2024 (WACV)
Reviewer , Association for Computational Linguistics, Rolling Review, 2024 (ACL)

Reviewer, International Conference on Learning Representations, 2024 (**ICLR**)

Reviewer, IEEE International Conference on Acoustics, Speech and Signal Processing, 2023 (**ICASSP**)

Reviewer, IEEE Transactions on Biometrics, Behavior, and Identity Science, 2023 (**TBiom**)

Reviewer, IEEE International Joint Conference on Biometrics, 2023 (**IJCB**)

Reviewer, IAPR International Journal on Document Analysis and Recognition, 2022 (**IJDAR**)

Reviewer, IEEE Multimedia, 2023 (**MM**)

Invited Lecture, Large Vision-Language Models, **Clarkson University**, 2023

President, Computer Science Graduate Student Association, 2021-22 (**CSE-GSA, UB**)

Writer, Towards Data Science (**TDS**)

Co-lead, Facebook Developer Circle Indore, 2017-19

Lead organizer, HackIndore, the largest hackathon of Central India, 2019

City Lead, Microsoft Student Partner, 2017-18

Founder and Head, SGSITS' first techno-learning club (#include), 2017-19

Head, Design, Entrepreneurship Cell SGSITS. 2017

Vice-Captain, Aeromodelling club SGSITS, 2016

Skills

Languages and Frameworks: **Python, Pytorch**, TensorFlow, JavaScript, Go (Golang), C++, C, SQL, HTML, CSS, R, Dart, Angular 8, React, Laravel, PHP, Pandas, Dask, PySpark, Numpy, React Native, Android, Java Spring

Databases & Cloud: SQL Server, MySQL, MongoDB, Redis, Cassandra, Solr, No-SQL, AWS, GCP

Software Tools: MATLAB, Visual Studio, Android Studio, Jupyter Notebook, Xcode